Self-Supervised Mapping and Localization by Predictive Learning

G. William Chapman
Sandia National Laboratories
Albuquerque, USA
gwchapm@sandia.gov

Andrew S. Alexander
University of California, Santa Barbara
Santa Barbara, USA
andyalexander@ucsb.edu

Frances S. Chance
Sandia National Laboratories
Albuquerque, USA
fschanc@sandia.gov

Michael E. Hasselmo
Boston University
Boston, USA
hasselmo@bu.edu

***Abstract*—Spatial navigation involves the formation of coherent representations of a map-like space, while simultaneously tracking current location in a primarily unsupervised manner. Despite a plethora of neurophysiological experiments revealing spatially-tuned neurons across the mammalian neocortex and subcortical structures, it remains unclear how such representations are acquired in the absence of explicit allocentric targets. Drawing upon the concept of predictive learning, we utilize a biologically plausible learning rule which utilizes sensory-driven observations with internally-driven expectations and learns through a contrastive manner to better predict sensory information. The local and online nature of this approach is ideal for deployment to neuromorphic hardware for edge-applications. We implement this learning rule in a network with the feedforward and feedback pathways known to be necessary for spatial navigation. After training, we find that the receptive fields of the modeled units resemble experimental findings, with allocentric and egocentric representations in the expected order along processing streams. These findings illustrate how a local and self-supervised learning method for predicting sensory information can extract latent structure from the environment.**

***Index Terms*—Spatial Navigation, Mapping and Localization, Neurally Inspired Computing**

I. Introduction

Behavioral navigation in allocentric, or map-like, space is essential for animals and robotic agents to navigate to goals which are outside of the immediate sensory space. While simple methods of spatial localization, such as path integration [1], can integrate self-motion signals to determine distance and direction traveled, such approaches are extremely sensitive to the presence of noise [2], would not persist across multiple experiences, and do not provide a full representation of space for more difficult navigation involving obstacles. In contrast, animal studies have shown that animals will mentally simulate planned trajectories in a combination of forward and backward search, similar to the A* algorithm used in artificial route planning [3], [4]. However, such planning requires that there first be an internal representation of location in space and that such a representation can be indexed by sensory information, while simultaneously translating relative goal-locations to egocentric operations. Here, we will develop a neurally-inspired algorithm to illustrate how future neuromorphic hardware may implement a robust low-powered, and unsupervised formation of allocentric coordinates.

Egocentric Representations: While view-invariant allocentric codes are required for longer-term navigation, animals rarely, if ever, receive direct allocentric information. Instead, the sensory information is encoded in the neocortex, where the location of stimuli will be relative to an animal, such as an object being distant and ahead in the visual field, or an object stimulating whiskers on the left side of an animal. Therefore, even if path integration alone is insufficient for formation of stable global maps, some degree of processing of sensory and motor information must occur, creating a transformation from the egocentric reference frame to a global map-like space. This transformation must also have an inverse transformation, so that allocentric codes can inform egocentrically-centered behavior, though such transformations may occur in other brain regions [5]. Neocortical regions are known to encode the location and identity of objects relative to an animal, beginning with primary sensory regions and convergent to association areas such as the parietal cortex. There have recently been numerous reports of neurons that respond to environmental boundaries in the postrhinal cortex [6] and retrosplenial cortex [7] as well as other egocentric responses, such as the relative location of a visual target [8] in posterior parietal cortex. The majority of these regions also receive afferent self-movement information and feedback information from the allocentric regions discussed below. Therefore, these regions are thought to be essential to transforming sensory information into allocentric codes.

Allocentric Representations: Extensive experimental evidence has shown that the medial temporal lobe (MTL) and associated structures in mammals contain allocentric codes for space. Place cells, found primarily in the hippocampus [9], respond to a single location (or a few locations) in an environment but will remap between environments, creating a reusable fabric for allocentric codes [10]. Grid cells in the medial entorhinal cortex [11] respond in a regular array of locations in the environment and project bidirectionally to place cells. They may form a basis for performing mental operations between place cell representations [4], [12]. Head direction cells that respond to the current compass direction of the head are found in many of these regions [13]. These MTL regions receive connections from the processed egocentric signals discussed above, but it is not clear how they directly

form from those signals

Learning Allocentric & Egocentric Predictions: Temporal prediction of spatial signals is known to create different latent representations depending on the specific input and target signals. Computational models that learn to encode future sensory states from current sensory states and motor information directly lead to allocentric representations, but do not form the same representations in an autoencoder approach without a predictive component [14]. These findings suggest that the prediction of sensory inputs extracts the transitions of sensory information in a reduced two-dimensional latent space [15], similar to generative machine learning models [16], suggesting that successor-representation learning is a general computation. Other work has shown that hidden units will form even more compressed representations in the form of regular grid-like activity [17] when tasked to predict future allocentric locations and given past allocentric locations and motor commands. More recent work has also shown that simple RNN models will exhibit a representational state in which principal components code separately for spatial and contextual information [18], similar to splitter cells [3] or remapping. However, allocentric predictions of the medial temporal lobe are also responsible for guiding learning in the neocortex, suggesting that stable map-like representations may guide sensory driven representations in uncertain environments [19]. These previous results show that in abstract networks temporal prediction of egocentric sensory signals gives rise to allocentric codes, and prediction of allocentric codes can lead to more compressed allocentric codes.

Current Approach: Learned Egocentric-Allocentric Associations: While the computational studies discussed above have shown the learning of allocentric spatial representations from temporal prediction of sensory representations, it remains unclear how such a representation would be learned in a biological substrate. Here we make use of the connectivity structure of a previous systems-level model of spatial cognition (the 'BBB models' below) [20], [21]. The BBB models hand-coded several functional cell types such as egocentric responses and place cells, and showed how they could be trained to drive each other in either a sensory-driven or memory-driven manner, providing a mechanism for allocentric mapping and egocentric movement. We utilize the same overall connectivity of this previous model, but instead of pre-determining the functional cell types and then training the weights, we implement a simplified model of learning. We replace each of the abstract regions from the BBB model with a biophysically-inspired model of canonical microcircuit, coupled with a learning rule which alters feedforward weights in order to predict future states [22].

II. Methods

Behavioral Task: We utilize a virtual agent moving in a 1m x 1m open field, sampled at 500Hz (dt=2ms) in order to give appropriate temporal resolution for the neural modeling described below. The agent moves according to a parametrically-driven approach described elsewhere [23], [24] with equations and parameters chosen (Table I) to realistically replicate the behavior of rats moving in an environment for 5 hours. Briefly, on each frame the rotational velocity is modeled as an Ornstein–Uhlenbeck (OU) process and running speed is modeled as a Rayleigh transformed OU process. The agent location and head direction are then generated by integrating these terms over time. Near boundary locations there was an anti-parallel repulsive force proportional to the distance from the wall, as well as a perpendicular force. Together, these two forces create a slowing down as the agent approaches a wall and causes a thigmotaxis-like following of the boundaries. We then utilize a three-dimensional rendering software [25] to generate agent-based visual scenes in colour and depth (RGB-D). In this rendering each wall is given a different colour in order to provide some grounding information as to the structure of the environment, and could be a stand in for sensory cues.

Variable	**Description** (Units)	**Range x Shape**
dt	Simulation timestep (s)	0.001
t	Time (s)	[0, 1200]
X	X location (m)	[0, 1]
Y	Y location (m)	[0, 1]
θ	Head direction (radians)	$[-\pi, \pi]$
S	Running Speed (m/s)	[0, 0.3]
ω	Rotational Velocity (radians/sec)	$[-\pi, \pi]$
I	Agent RGB View (1)	[0, 1] x [60, 80, 3]
D	Agent Depth View (m)	[0, 1.4] x [60, 80]

TABLE I
BEHAVIORAL VARIABLES FOR SIMULATION

A. Neural Model

Macroscale Connectivity: Rather than typical feedforward layers of machine learning approaches, we utilize a restricted cyclic connectivity between modules, which reflects the macroscale connections among brain regions strongly implicated in spatial navigation of rodents (Fig. 1) [21]. External stimuli come from two major sources, the egocentric sensory (vision) and the motor/proprioceptive pathway. The vision pathway begins with the identity and distance of the visual field array projecting to a visual modules termed 'ventral' and 'dorsal' respectively. Similarly, the linear velocity and rotational velocity from the behavioral simulation project to a singular 'action' module. The two visual streams and the action module have convergent feedforward pathways onto the retrosplenial cortex (RSC), which in turn projects to the perirhinal cortex (PR), and hippocampus (HPC), representing the primary feedforward visual to allocentric pathway. The entorhinal module (MEC) also receives direct feedforward activity from the 'action' module and projects to the feedforward compartment of the hippocampal module. Feedback connections travel in the opposite order and along the same route. Thus, the dorsal, ventral, and action modules are the only ones which directly relate to the environment and provide a basis for interpretation later.

Microcircuit Connectivity: The model that we utilize here relies on separate feedforward and feedback pathways, with dense and normally distributed weights between individual

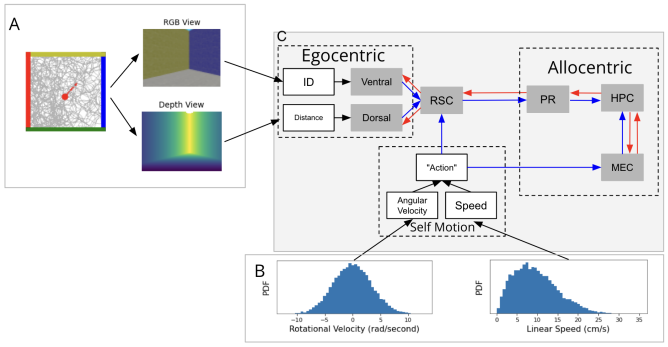

Fig. 1. Overall simulation paradigm. **A** (Left) As an agent (red dot, facing towards arrow) navigates around an environment (grey line), behavioral frames are generated. (Top) RGB images are encoded with a one-hot paradigm to provide object identity. (Bottom) Depth images are also generated, representing early processing along the ventral visual stream. **B** Behaviorally realistic rotational velocity (left) and running speed (right) distributions are used to generate the overall trajectory from A. These variables are separately encoded on each frame into a self-motion representation. **C** Shows the macroscale connectivity between modules. Black lines indicate static weights which encode information from the behavioral simulation. Blue lines indicate inter-area feedforward synapses which target onto the granular layer of neocortex or dentate gyrus populations and carry overall activation. Red lines indicate feedback weights which originate from the infragranular layers of neocortex or the CA1 region of hippocampus and terminate on distal dendrites, carrying burst rates.

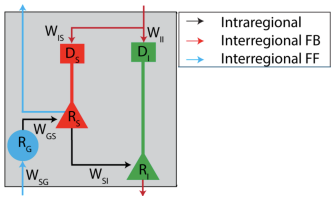

Fig. 2. Architecture of the predictive module: The structure of each module, where rates ('R') of each submodule represent the firing rate of a given population of neurons, and weights are labeled according to their source and target submodules. Units are coloured by their laminar location, also denoted by subscript (G)ranular, (S)uperficial, (I)nfragranular, and (D)istal. Weights are coloured by according to their primary function in the feedforward (cyan), feedback (maroon) or local (black) pathways.

units that are connected by the macroscale constraints above. Large-scale summaries of cortical microcircuit connectivity [26], [27] suggests that in neocortex there is a primarily three-layer separation of granular (G), superficial (S), and infragranular (I) layers (see Figure 2). The granular layer receives feedforward inputs to point-like neurons, while the superficial and infraganular layers contain pyramidal neurons which are modeled with two-compartments. Pyramidal neurons receive feedforward information into their somatic compartments, and feedback information into the distal dendrites (D). In hippocampus, these modules map to to dentate gyrus, CA3, and CA1 respectively, and the feedforward activity travels through the Schaffer Collaterals [28], and feedback through the temporammonic pathway and input to pyramidal distal dendrites from the entorhinal cortex [29], [30]. Thus, the overall connectivity and in the hippocampal module is the same as within the neocortex.

Following prior work, granular neurons follow the general recurrent neural network (RNN) formulation reflecting their compact physiology. For each Layer "L":

$$\begin{aligned} \tau \frac{dv_L}{dt} &= -v_L + \sum_{N \in A} W_{NL} r_N \\ r_L(t) &= \sigma(v_L(t)) \end{aligned} \tag{1}$$

Where τ is the population specific time constant (10ms throughout), v represents the membrane potential. W represents the weights of afferent connections from presynaptic populations A (in the summation index), including recurrent connections among a population. For the case of the granular neurons, A is the set of all external inputs and feedforward afferents (black lines in Figure 1). σ is the nonlinear activation function converting membrane potential v to action potential r, and we utilize the hyperbolic tangent for this activation function.

The pyramidal units contain the same leaky-integrator format for their somatic potential, but incorporate an additional term representing the potential of the distal dendrites and overall burst probability:

$$\begin{aligned} \tau \frac{dv_L}{dt} &= -v_L + \sigma(u_L/D_d) + \sum_{N \in A} W_{NL} r_N \\ r_L &= \sigma(v_L) \\ \tau_u \frac{du_L}{dt} &= -u_L + 2\frac{dr_L}{dt} + \sum_{N \in B} W_{Nu} b_N \\ p &= \sigma(u) \\ b_L &= r_L \odot p_L \end{aligned} \tag{2}$$

The additional term u represents the distal dendritic compartment, and the summation index B represents the population of neurons projecting to that compartment, compared to those projecting to the soma (summation index A). For superficial units A is the recurrent connections and activity of the local granular units, and for infragranular units represents the recurrent connections and activity of the local superficial units. For both sets of pyramidal neurons, B is the set of feedback connections (red lines in Figure 1). 'p' represents the probability of a burst event and follows a sigmoidal function of the distal dendrites [31]. 'b' represents the burst rate of the pyramidal neurons and is the direct product of the overall firing rate and the burst probability. The middle terms in both the somatic potential and dendritic potential equations represent subthreshold coupling and backpropagating action potentials (BAP) respectively, and allow mixing of feedforward and feedback signals, and increasing the burst rate (b) when these inputs are coincident in time.

B. Learning Rule

We implement continuous online 'training' through the use of a learning rule [32] on the feedforward synapses of pyramidal units. This rule is applied on each timestep of training, without a global target signal, explicit error propagation, or typical machine-learning modifications such as learning-rate schedules or moment-averaging [33]. The learning rule

combines a classical "pre-post" term with a top-down gain based on the difference in instantaneous and long term burst probabilities:

$$\Delta W_{XY} = \eta((p_Y - \bar{p_Y}) \odot r_Y) \otimes r_X \tag{3}$$

Where X is the presynaptic population and Y is the postsynaptic, and $\bar{p_Y}$ represents the low-pass burst probability on the temporal order of behavioral variables. The learning rate η is held at 10^{-4} throughout.

In the case of a smoothly varying input the difference between instantaneous burst probability and low pass average is approximately the temporal derivative. Then, letting x and y denote individual units for notational simplicity, and following the chain rule for the derivative of p_Y:

$$\begin{aligned}
\Delta W_{xy} &= \eta((p_y - \bar{p_y})r_y)r_x \\
&= (\eta r_x r_y)(p_y - \bar{p_y}) \\
&\approx (\eta r_x r_y)\frac{dp}{dt} \\
&\approx (\eta r_x r_y)\frac{d\sigma(u)}{du}\frac{du}{dt} \\
&\approx (\eta r_x r_y)(\sigma(u)(1-\sigma(u))\frac{1}{\tau_u}(-u + 2\frac{dr_y}{dt} + \sum Wb)
\end{aligned} \tag{4}$$

Evaluating at a given value of u:

$$\Delta W_{xy}(u) \propto (\eta r_x r_y)\frac{dr_y}{dt}F(u) \tag{5}$$

From this formulation the instantaneous change in feedforward synaptic weight is proportional to the presynaptic activity times the postsynaptic activity (Hebbian associative learning) times the instantaneous change in unit rate (temporal error) and multiplied by the distal dendritic potential (feedback gain).

C. Analyses

Generalized Linear Model: When attempting to classify the receptive fields of units we utilize a general framework:

$$\begin{aligned}
R(t) = \sigma(&\beta_0 + \beta_{pc}\lambda_{PC} + \beta_{hd}\lambda_{HD} + \\
&\beta_{ls}\lambda_{LS} + \beta_{ebc}\lambda_{EBC} + \beta_{grid}\lambda_{grid})
\end{aligned} \tag{6}$$

Where $R(t)$ is the activation of a given unit at a single point in time and σ is the nonlinear activation function. In this equation, PC=place cell, HD=head direction cell, LS=running speed cell, EBC=egocentric boundary cell and grid=grid cell. β_0 Is the baseline bias in activity, and the remaining β terms are strengths of behavioral variable influence on activation. The λ terms represent parametric tuning curves that match the general shapes of experimental findings and are individually optimized for each unit:

$$\begin{aligned}
\lambda_{pc} &= e^{-\frac{1}{2(1-\rho^2)}((\frac{x-\mu_x}{\sigma_x})^2 - 2\rho(\frac{x-\mu_x}{\sigma_x}\frac{y-\mu_y}{\sigma_y}) + (\frac{y-\mu_y}{\sigma_y})^2)} \\
\lambda_{hd} &= (\cos(\theta - \mu_{hd}))^k \\
\lambda_{ls} &= \lfloor \beta_s s + (1-\beta_s)e^{-\frac{(s-\mu_s)^2}{2\sigma_s^2}} \rfloor \\
\lambda_{ebc} &= max(I \odot \lfloor_{l\in L, r\in R} e^{-((\frac{l-\mu_l}{\sigma_l})^2) + (\frac{r-\mu_r}{\sigma_r})^2)}) \\
\lambda_{grid} &= \frac{1}{3}\lfloor \sum_{n=0}^{2} \cos(2\pi \frac{xe_{\theta+n\pi/3}}{\Lambda} + \phi) \rfloor
\end{aligned} \tag{7}$$

Within each subequation behavioral variables are as in table I and valid ranges for tuning parameters and their description is in table II. Classification of place cells used a two dimensional Gaussian centered at any point in the environment. Classification of head direction used a cosine function with mean μ_{hd} and tuning parameter k which decreases with width of the tuning curve. Running speed tuning is modeled as the sum of a linear slope and a superimposed Gaussian function, enabling a diverse range of speed-curves including linear increase, saturating responses, and maximum response models [8]. Classification of egocentric boundary cell tuning used a two dimensional Gaussian displaced from the current location of the agent by μ_l and μ_r relative to the direction of the animal. In equation 7, I is an indicator function evaluated at 0.1cm increments in the x and y direction relative to the animal for the presence of a boundary, and then rotated to the current movement direction. The second term of the EBC response is a 2D Gaussian intensity function evaluated at Rostral-Caudal (R) and Lateral-Medial (L) offsets from the agent location. The grid cell model is the sum of three spatial sinusoids with frequency Λ, spaced 60° apart and overall orientation to the environment θ [34]. Finally, we do not explicitly provide an equation for boundary vector cells (BVCs), since a place field with unconstrained location and covariance is degenerate with the typical equations used to model BVCs. In the results, however, when a unit is significantly modulated by the "place cell" predictor but is elongated along a boundary we identify it as a BVC instead of a place cell.

Parameter	Description	Range
$\sigma_{x/y}$	X/Y width of place cell (cm)	
ρ	X/Y Covariance (angle) of place cell (1)	
$\mu_{x/y}$	X/Y center of place cell (cm)	
k	Tuning width of HD cell (1)	[1/3, 3]
μ_{hd}	Center of HD cell (rad)	[$-\pi$, π]
β_{s_1}	Slope of speed cell (s/cm)	[0, 30]
β_{s_2}	Strength of Gaussian speed cell (s/cm)	[0, 30]
L	Lateral offset from the agent (cm)	[-50, 50]
R	Rostral offset from the agent (cm)	[-50, 50]
$\mu_{R/L}$	R/L center of egocentric boundary cell (cm)	[-50, 50]
θ	Angle offset of grid cells (rad)	[0, $\pi/3$]
ϕ	Phase offset of grid cells (rad)	[$-\pi$, π]
Λ	Spatial scale of grid cells (cm)	[20, 50]

TABLE II
TUNING CURVE PARAMETERS FOR EVALUATING FUNCTIONAL CELL TYPES

Measuring Representation Learning: After every simulated three minute interval ("epoch") we evaluate the tuning of all learned units. This is done by, without enabling learning,

simulating the network on a separately generated 60 minute session and then down-sampling the behavior and unit activity to 10Hz (36K frames). The activity from this "test" session is then fit to the explicit tuning curve given in equation 6 by gradient descent (SGD, minimum error over 10 random initializations of 5-fold verification), as well as submodels that each drop one of the predictors. β values were constrained to be non-negative during the fitting process. For each submodel we compare the mean-square error in predicted activity compared to the true activity and generate an F-statistic, which is analogous to the log-likelihood approaches from previous approaches [8]. The proportion of units that are statistically significantly tuned ($p < 0.001$) is measured on each epoch and is reported below. In Figures 4 and 6 all tuning curves are represented as normalized activity in 2D (spatial), polar (head direction), or 1D (running speed) bins.

III. Results

A. Learned Responses

We first investigate the final somatic event rate tuning curves and burst-rate tuning curves, as summarized in Figure 4. Speed tuning was present in all regions, with a total of 87% (890/1024) units being significantly modulated by running speed. Consistent with expected findings, the feedforward signals of the RSC module contains a number (80/128, 63%) of units which respond to the egocentric boundary predictor. These can be recognized in Figure 4 by the strong responses around the edges of the environment, which were accompanied by head direction selectivity for different boundaries. Another 41 (32%) were classified as head-direction sensitive, though notably no units had significant tuning to both head direction and EBC tuning. In the feedback layer of RSC, one unit (unit 0 in Figure 4) showed significant EBC response. Half (64/128) of FB RSC units responded to head direction, and 108 (84%) to movement speed. Superficial perirhinal units expressed several BVC-like responses, identified as a response to the place cell predictor, but with a variance along one axis at least three times greater than along the other (52/128 41% BVC, 23/128 18% PC). These were largely conjunctive (67/75, 90%) with head direction tuning. Deep PR units generally expressed the same variables as their superficial counterparts, but with a more pronounced tendency for place cell (48/128, 38%) than BVC (12/128, 9%) response. Hippocampal feedforward units were generally place-cell like (115/128, 89%), with the remaining portion being non-spatially responsive. The majority of units were conjunctive with head direction (102/128, 80%). Deep hippocampal units were tuned to both place (27/128, 21%) and BVC (47/128, 37%), with a slightly less tendency to be head direction tuned (45/128, 35%). Two deep hippocampal units were significantly tuned to the grid predictor, but upon visual inspection primarily tended to fire at three corners of the environment. Entorhinal units showed a combination of speed (108/125, 84%) and head direction (42/128, 33%) tuning. A subset of units (18/128, 14%) showed tuning to location and exhibited a mixture of near-boundary and near-center responses.

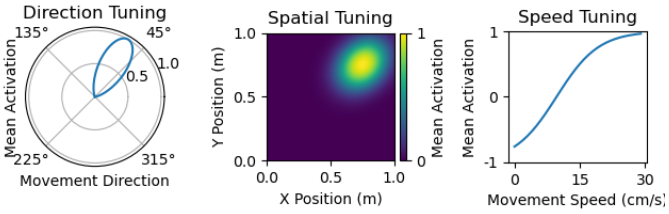

Fig. 3. To compare learned representations to neuroscience literature, we visualize tuning curves in response to behavioral variables, with idealized responses shown here. **Left** Movement direction is illustrated on the polar axis, while mean activation is on the radial. An idealized movement-tuned cell is shown, which responds to movement to the north-east. **Middle** Spatial location is illustrated on the 1m x 1m arena, with an idealized place-cell illustrated. **Right** Movement speed, with an idealized neuron which responds only to increased speed.

Order of Learning: We also evaluated the rate at which representations converged to their final value over the course of training. On each epoch we found the mean correlation coefficient between the holdout session activity at that point in training and the holdout activity at the end of training (Figure 5). Contrary to expectations, it was neither the most (RSC) or least (HPC) driven representations which converged most quickly. Instead the perirhinal, which was predominantly responsible for encoding distance to a singular boundary, converged within the first tens of epochs. This region was then followed by hippocampus and retrosplenial, suggesting that the feedforward weights of perirhinal aligned to their inputs before the hippocampal module received a consistent enough signal to create place-like representations. Feedback bursting signals converged more slowly than any of the feedforward activities (Figure 5).

B. Response to Novel Environment

We now consider a case where the feedforward activity violates the trained expectations. At the end of training we introduce an additional session in which an unexpected barrier with a unique identity encoding has been inserted into the middle of the environment (y=0.48 - 0.52, x = 0.25 - 0.75). For clarity of analysis, we disable learning during these experiments, and measure mean responses across a simulated 60 minute exposure. The tuning responses of units in this condition are demonstrated in Figure 6, and fitting of the curves is performed as in the baseline condition, and the chosen example units are the exact ones shown in Figure 4.

In general, individual feedforward allocentric tuning curves were disrupted in the novel environment, while speed tuning and EBC responses were consistent. In superficial RSC, 75 (93%) of the units which were EBC tuned in the baseline continued to be so in the novel environment, while only 10 (24%) remained modulated by head direction. In superficial PR all units previously classified as BVCs continued to be so, but their preferred boundary and x/y standard deviations changed. Hippocampal place-cell tuning was only present in a fraction of the units from baseline (2/115, 2%). Feedback activity was largely disrupted as much as the corresponding feedforward

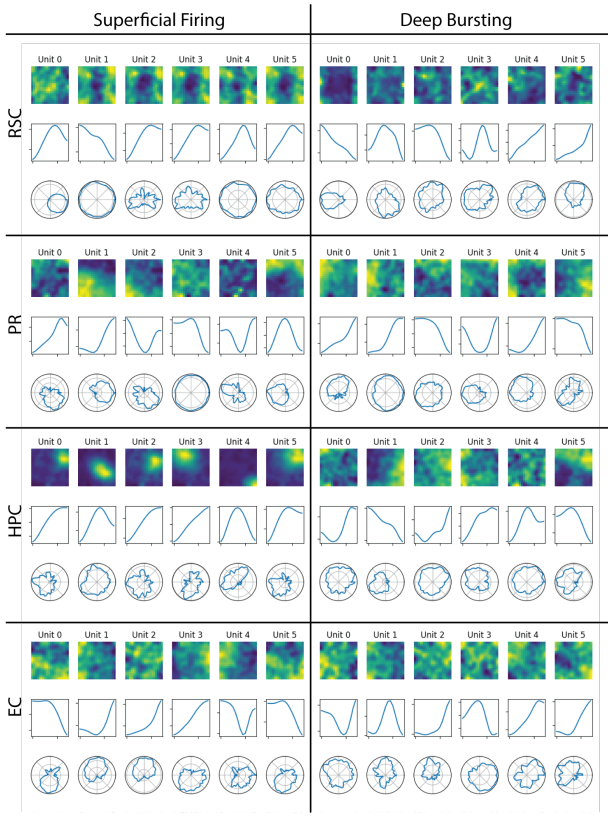

Fig. 4. For each box the feedforward (left) or feedback (right) activity is shown for six randomly selected units at the end of training, utilizing the visualization of Figure 3. **RSC-FF** Many units exhibit egocentric-boundary responses, firing near the boundaries of the environment, regardless of head direction (unit 3). Others respond to head direction (unit 0). Nearly all respond to movement speed. **RSC-FB** Compared to FF, there is relatively little spatial modulation, though head direction and speed tuning remain prominent. **PR-FF** Spatial tuning (top) tended to be strongly locked to one boundary (unit 5) or corner (unit 1). Head direction and speed tuning are present. **PR-FB** Tuning is similar to the FF case. **HPC-FF** (Top) Place-cell like representations are found, with centers distributed throughout the environment. (Bottom) A subset of units are conjunctive with head direction (unit 0), while others are not (unit 1). **HPC-FB** Spatial representation is more PR-like, with receptive fields on the boundaries of the environment and higher dispersion. **EC-FF** Spatial modulation exists for a subset of units, and tends to be anchored to an environmental boundary and conjunctive with head direction. **EC-FB** Feedback spatial activity is dispersed, but head direction tuning is still present.

activity, with the exception of deep RSC where head direction tuning was preserved in 93% (60/64) of the original units. Of the feedback PR activity which was tuned to place-cell predictors in the familiar environment, a significant portion changed to being BVC tuned (20/48, 42%). Across all regions, feedforward and feedback, speed tuning was maintained in 793 of the original 890 (89%).

While Figure 6 clearly shows that individual unit representations are disrupted in the novel environment, we next asked whether it was necessarily the case that the information coded by distributed representations was disrupted. To address this, for each region and each behavioral variables (location, head direction, and speed) we train a linear decoder to produce the minimum mean-square error readout both during the familiar and novel environment. For head direction, the decoder was

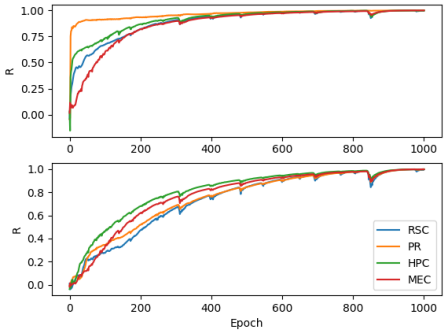

Fig. 5. Correlation coefficient of feedforward (top) and feedback (bottom) representational similarity on each epoch compared to at the end of training. Feedforward activity stabilizes within the first 200 epochs, while feedback activity undergoes changes further into the training period.

trained on the mean sum error of sin and cosine components of the angle. This was done using increments of ten units and five repetitions of 5-fold cross validation. Overall, we find that in the baseline (familiar) condition this distributed readout was able to decode location from all regions except RSC, head direction from PR and RSC, and movement speed from all regions (Figure 7, Top). In contrast, position and head direction information was not significantly represented, even in a distributed manner, in any regions during the novel environment (Figure 7). Running speed was still coded in all regions, only slightly less prominently than in the familiar session.

IV. Discussion

The above results illustrate that a self-supervised hierarchical system imbued with a predictive learning rule can learn latent representations of an underlying behavioral space. In the current implementation, sensory information is carried by a feedforward signal while feedback is carried by a separate value. However, previous work has illustrated that these two signals can be multiplexed into firing rate and burst rate in models with higher biological fidelity [22]. By greedily minimizing the difference between the feedforward and feedback signals, the learning rule helps to guide a synapse update to minimize the moment-to-moment changes in somatic potential. After simulating this system in a passive navigation setting for an extended period, we see the emergence of internal representations which resemble those found in animal studies. In particular we find the emergence of units similar to egocentric boundary cells [35] which respond to the presence of a barrier based on its egocentric displacement relative to the current direction and position of the agent. As they receive a combination of sensory distance signal (‘dorsal’), self-motion information (‘action’), and generate an allocentric location signal (‘PR’ boundary cells), we hypothesize that these implement a form of transformation from egocentric to allocentric signal [36]. The next region in the hierarchy then provides a response based on distance from environmental

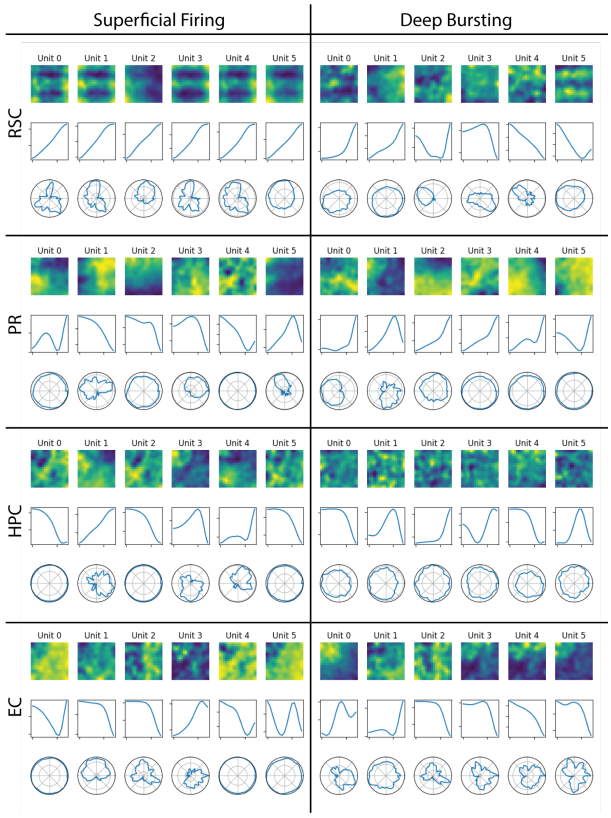

Fig. 6. *Novel Environment Tuning* **RSC** Many units (eg: FF unit 1) continue to exhibit EBC-like tuning, showing a strong response near the familiar and novel boundary. Head direction tuning is largely disrupted. **PR-FF** A number of PR units continue to show HD tuning. Many spatial tunings have shifted compared to the baseline, but still tend to fire near an edge of the outer boundary. **PR-FB** Deep PR activations are more likely to be tuned to BVC responses than during baseline, but these responses tend to be wide. **HPC** Place-cell like representations are largely disrupted. Units that remain HD tuned tend to have a different preferred direction. **EC** Some units remain tuned to head direction, but spatial modulation is largely gone.

boundary signal (an allocentric boundary response) that the final region ('HPC') integrates over time to provide allocentric information even when far from the boundary locations. This place-centric code can then provide additional contextual information to the perirhinal to resolve any ambiguity in feedforward signal, such as two barriers with similar appearance. The general internal representations we find are in line with both experimental findings and theoretical [20] expectations. In particular, consistent with experimental data, the model shows formation of neural responses that progress from regions showing egocentric representation of barriers (egocentric boundary cells in retrosplenial) to allocentric representation of barriers (boundary vector cells) to allocentric representation of the spatial location of the agent (place cells in hippocampus).

An interesting comparison can be drawn between the responses learned here and those learned in Recanatesi et. al 2021 [14], where a single layer RNN trained in a similar manner produces place-like representations but not the types of responses we see in the intermediate RSC and PR modules. This suggests that while a single RNN will tend to extract

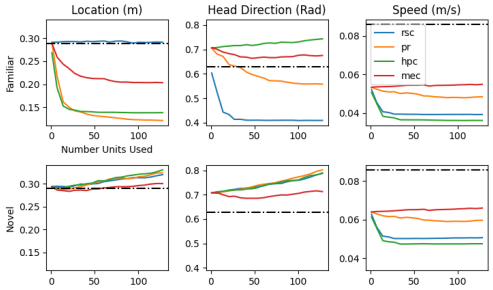

Fig. 7. *Behavioral Decoding* Mean-square error of linear readout of behavioral variables in each modeled region during foraging in the familiar environment. Dashed black line indicates chance. **Top** Location can be significantly read out in all regions except RSC. Conversely, head direction is most strongly *linearly* read out from RSC and PR. Speed is significantly represented everywhere, but is more redundant in MEC and PR than RSC and HPC. **Bottom** In the novel environment with inserted barrier, the decoding is greatly altered. Using the same error scales as the baseline, location and head direction are not significantly read out from any regions. Speed continues to be represented in all regions, with a slight decrease in performance compared to the familiar environment.

direct successor-representation-like activity, the restriction of weights by additional hierarchy leads to intermediate representations that facilitate translation between the sensory and successor representations.

Implications for Neuromorphic: Previous work has shown how neurally-inspired approaches can provide high fidelity and low-power implementations of simultaneous localization and mapping by combining path integration with reset mechanisms when encountering a familiar landmark [37], [38]. The current work expands upon such implementations by allowing a more continuous and graded interaction of these two pathways, while learning in an online and biologically inspired manner. The predictive learning rule, by virtue of interactions only by pre-post interactions and potential implementation by spike-burst multiplexing has the potential to be applied to spiking neuromorphic hardware [39]. While the current approach utilizes real-valued inputs, within the feedforward and feedback pathway, equation 4 and 5 show the equivalence of the rule used here and our previous spiking based implementations [22] when combined with short-term plasticity effects. The online learning may indeed be necessary for flexible behavior, as the results of the novel environment experiment (Figure 6) showed that changes in sensory inputs can greatly diminish allocentric codes when learning is disabled. While allocentric codes do not spontaneously adjust to the novel environment in our model, experimental evidence suggests that remapping may occur within the first few minutes of exposure to a novel environment [40], but only if learning is enabled on-device. Overall, the work here can be modified to be spiking, relies on locally dense but globally sparse connectivity, and utilizes an online learning rule, all of which lend to deployment to neuromorphic hardware.

Summary: The simulation presented here demonstrates that a biologically-based model of learning can generate responses of individual neurons corresponding to functional

cell types in neurophysiological experiments. This includes a hierarchical shift from sensory-driven to memory-driven fields, as regions shift further from sensory inputs. The hierarchical integration of feedforward and feedback information into the predictive learning rule appears to be essential for this transition. Future work will implement the spiking version of this approach on neuromorphic platforms, and utilize the learned representations for goal-driven navigation in continually changing environments.

REFERENCES

[1] B. L. McNaughton, F. P. Battaglia, O. Jensen, E. I. Moser, and M.-B. Moser, "Path integration and the neural basis of the 'Cognitive map'," *Nature Reviews Neuroscience*, p. 16, 2006.

[2] K. Hardcastle, S. Ganguli, and L. M. Giocomo, "Environmental Boundaries as an Error Correction Mechanism for Grid Cells," *Neuron*, vol. 86, no. 3, pp. 827–839, 2015.

[3] J. Ferbinteanu and M. L. Shapiro, "Prospective and Retrospective Memory Coding in the Hippocampus," *Neuron*, vol. 40, no. 6, pp. 1227–1239, Dec. 2003.

[4] U. M. Erdem and M. E. Hasselmo, "A goal-directed spatial navigation model using forward trajectory planning based on grid cells," *European Journal of Neuroscience*, vol. 35, no. 6, pp. 916–931, 2012.

[5] J. R. Hinman, G. W. Chapman, and M. E. Hasselmo, "Neuronal representation of environmental boundaries in egocentric coordinates," *Nature Communications*, vol. 10, no. 1, p. 2772, Dec. 2019.

[6] P. A. LaChance, T. P. Todd, and J. S. Taube, "A sense of space in postrhinal cortex," *Science Advances*, p. 12, 2019.

[7] A. S. Alexander, L. C. Carstensen, J. R. Hinman, F. Raudies, G. W. Chapman, and M. E. Hasselmo, "Egocentric boundary vector tuning of the retrosplenial cortex," *Science Advances*, Jul. 2019.

[8] A. S. Alexander, J. C. Tung, G. W. Chapman, A. M. Conner, L. E. Shelley, M. E. Hasselmo, and D. A. Nitz, "Adaptive integration of self-motion and goals in posterior parietal cortex," *Cell reports*, vol. 38, no. 10, p. 110504, 2022.

[9] J. O'Keefe and J. Dostrovsky, "The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat," *Brain Research*, vol. 34, no. 1, pp. 171–175, Nov. 1971.

[10] R. Muller and J. Kubie, "The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells," *The Journal of Neuroscience*, vol. 7, no. 7, pp. 1951–1968, Jul. 1987.

[11] L. L. Colgin, T. Denninger, M. Fyhn, T. Hafting, T. Bonnevie, O. Jensen, M.-B. Moser, and E. I. Moser, "Frequency of gamma oscillations routes flow of information in the hippocampus," *Nature*, vol. 462, no. 7271, pp. 353–357, Nov. 2009.

[12] D. Bush and N. Burgess, "A hybrid oscillatory Interference/Continuous attractor network model of grid cell firing." *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 34, no. 14, pp. 5065–79, Apr. 2014.

[13] J. S. Taube, "The head direction signal: Origins and sensory-motor integration." *Annual review of neuroscience*, vol. 30, pp. 181–207, Jan. 2007.

[14] S. Recanatesi, M. Farrell, G. Lajoie, S. Deneve, M. Rigotti, and E. Shea-Brown, "Predictive learning as a network mechanism for extracting low-dimensional latent space representations," *Nature communications*, vol. 12, no. 1, p. 1417, 2021.

[15] K. L. Stachenfeld, M. M. Botvinick, and S. J. Gershman, "The hippocampus as a predictive map," *Nature Neuroscience*, vol. 20, no. 11, pp. 1643–1653, 2017.

[16] K. Li, A. K. Hopkins, D. Bau, F. Viégas, H. Pfister, and M. Wattenberg, "Emergent world representations: Exploring a sequence model trained on a synthetic task," *arXiv preprint arXiv:2210.13382*, 2022.

[17] A. Banino, C. Barry, B. Uria, C. Blundell, T. Lillicrap, P. Mirowski, A. Pritzel, M. J. Chadwick, T. Degris, J. Modayil, G. Wayne, H. Soyer, F. Viola, B. Zhang, R. Goroshin, N. Rabinowitz, R. Pascanu, C. Beattie, S. Petersen, A. Sadik, S. Gaffney, H. King, K. Kavukcuoglu, D. Hassabis, R. Hadsell, and D. Kumaran, "Vector-based navigation using grid-like representations in artificial agents," *Nature*, vol. 557, no. 7705, pp. 429–433, 2018.

[18] I. I. Low, L. M. Giocomo, and A. H. Williams, "Remapping in a recurrent neural network model of navigation and context inference," *Elife*, vol. 12, p. RP86943, 2023.

[19] G. Doron, J. N. Shin, N. Takahashi, C. Bocklisch, S. Skenderi, M. Drüke, L. de Mont, M. Toumazo, M. von Heimendahl, M. Brecht, R. Naud, and M. E. Larkum, "Perirhinal input to neocortical layer 1 controls learning," Neuroscience, Preprint, Jul. 2019.

[20] A. Bicanski and N. Burgess, "A Neural Level Model of Spatial Memory and Imagery," *eLife*, vol. 7, e33752, 2018.

[21] P. Byrne, S. Becker, and N. Burgess, "Remembering the Past and Imagining the Future: A Neural Model of Spatial Memory and Imagery," *Psychological Review*, p. 36, 2007.

[22] G. W. Chapman and M. E. Hasselmo, "Predictive learning by a burst-dependent learning rule," *Neurobiology of Learning and Memory*, p. 107826, 2023.

[23] F. Raudies and M. E. Hasselmo, "Modeling Boundary Vector Cell Firing Given Optic Flow as a Cue," *PLoS Computational Biology*, vol. 8, no. 6, pp. 1–17, 2012.

[24] T. M. George, M. Rastogi, W. de Cothi, C. Clopath, K. Stachenfeld, and C. Barry, "RatInABox, a toolkit for modelling locomotion and neuronal activity in continuous environments," *Elife*, vol. 13, p. e85274, 2024.

[25] M. Chevalier-Boisvert, "Miniworld: Minimalistic 3D Environment for RL & Robotics Research," 2018.

[26] D. J. Felleman and D. C. Van Essen, "Distributed hierarchical processing in the primate cerebral cortex." *Cerebral cortex*, vol. 1, no. 1, pp. 1–47, 1991.

[27] S. Haeusler and W. Maass, "A statistical analysis of information-processing properties of lamina-specific cortical microcircuit models," *Cerebral Cortex*, vol. 17, no. 1, pp. 149–162, 2007.

[28] S. O'Mara, "The subiculum: What it does, what it might do, and what neuroanatomy has yet to tell us," *Journal of Anatomy*, vol. 207, no. 3, pp. 271–282, Sep. 2005.

[29] H. Takahashi and J. C. Magee, "Pathway Interactions and Synaptic Plasticity in the Dendritic Tuft Regions of CA1 Pyramidal Neurons," *Neuron*, vol. 62, no. 1, pp. 102–111, Apr. 2009.

[30] Z. Friedenberger, E. Harkin, K. Tóth, and R. Naud, "Silences, spikes and bursts: Three-part knot of the neural code," *The Journal of Physiology*, vol. 601, no. 23, pp. 5165–5193, 2023.

[31] R. Naud and H. Sprekeler, "Sparse bursts optimize information transmission in a multiplexed neural code," *Proceedings of the National Academy of Sciences*, vol. 115, no. 27, pp. E6329–E6338, Jul. 2018.

[32] A. Payeur, J. Guerguiev, F. Zenke, B. A. Richards, and R. Naud, "Burst-dependent synaptic plasticity can coordinate learning in hierarchical circuits," *Nature Neuroscience*, vol. 24, no. 7, pp. 1010–1019, Jul. 2021.

[33] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," Jan. 2017, comment: Published as a conference paper at the 3rd International Conference for Learning Representations, San Diego, 2015.

[34] N. Burgess, C. Barry, J. O. Keefe, B. E. T. Al, and J. O'Keefe, "An oscillatory interference model of grid cell firing," *Hippocampus*, vol. 000, pp. 1–3, 2007.

[35] A. S. Alexander, J. C. Robinson, H. Dannenberg, N. R. Kinsky, S. J. Levy, W. Mau, G. W. Chapman, D. W. Sullivan, and M. E. Hasselmo, "Neurophysiological coding of space and time in the hippocampus, entorhinal cortex, and retrosplenial cortex," *Brain and Neuroscience Advances*, vol. 4, p. 239821282097287, Jan. 2020.

[36] A. A. Wilber, B. J. Clark, T. C. Forster, M. Tatsuno, and B. L. McNaughton, "Interaction of Egocentric and World-Centered Reference Frames in the Rat Posterior Parietal Cortex," *Journal of Neuroscience*, vol. 34, no. 16, pp. 5431–5446, 2014.

[37] M. J. Milford, G. F. Wyeth, and D. Prasser, "RatSLAM: A hippocampal model for simultaneous localization and mapping," in *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004*, vol. 1. IEEE, 2004, pp. 403–408.

[38] B. Komer, P. Jaworski, S. Harbour, C. Eliasmith, and T. DeWolf, "BatSLAM: Neuromorphic Spatial Reasoning in 3D Environments," in *2022 IEEE/AIAA 41st Digital Avionics Systems Conference (DASC)*, Sep. 2022, pp. 1–8.

[39] M. Stuck and R. Naud, "Burstprop for Learning in Spiking Neuromorphic Hardware."

[40] D. Lee, B.-J. Lin, and A. K. Lee, "Hippocampal place fields emerge upon single-cell manipulation of excitability during behavior." *Science (New York, N.Y.)*, vol. 337, no. 6096, pp. 849–853, Aug. 2012.